Journal of Computational and Applied Mathematics 11 (1984) 251–258
North-Holland

Rectified approximations for the solution of nonlinear equations

Shlomo BREUER and Gideon ZWAS
Department of Applied Mathematics, Tel-Aviv University, Tel-Aviv, Israel

Received 9 April 1984
Revised 6 June 1984

Abstract: In this paper we introduce the 'rectification method' for the construction of algorithms of pre-designed order r for the solution of nonlinear equations $f(x)=0$. Our method is based upon the derivation of a rectified approximation $g(x)$ to $f(x)$, via Padé formulas, such that the application of the Newton–Raphson iterations to g generates the desired rth order algorithm. Various properties of g are explored as are recursive relations among rectified approximations associated with successive orders of convergence. It is demonstrated that the use of g in favor of f can relax standard sufficient conditions assuring convergence of the iterations.

Keywords: Rectified approximations, rootfinding, iterations of rth order.

1. Introduction

In this paper we consider iterative procedures for the computation of roots of a sufficiently smooth function $f(x)$. To begin with, we shall treat the case of a simple root at $x=\rho$, to be approximated by iterations $x_{n+1}=\Phi(x_n)$ of a *pre-designed* order r, i.e., $|x_{n+1}-\rho|=O(|x_n-\rho|^r)$.

We propose to arrive at the rth order iterative formula through the application of the Newton–Raphson method to an associated 'rectified' function $g(x)$, instead of $f(x)$. The function $g(x)$, like $f(x)$, will be shown to possess a simple root at $x=\rho$ and to satisfy the additional conditions $g'(\rho)=f'(\rho)$, $g^{(k)}(\rho)=0$ for $k=2, 3, \ldots, r-1$. Thus $g(x)$ has an $(r-1)$st order contact with the tangent line at $x=\rho$ and is hence termed the rectified function. The *final* iterative formulas we obtain for specific values of r are not novel (see halley [4] for $r=3$, Kiss [8] for $r=4$, 5 and Traub [9, p. 88 ff.] for higher values of r). We believe, however, that our method produces insight into numerical root approximation in that it links higher order root-finding algorithms to the rectified function $g(x)$.

Once g has been constructed, the extensive available theory of Newton–Raphson iterations can be applied. Moreover, in addition to generating rth order iterations the method enables us, in many cases, to relax routine sufficient conditions assuring convergence.

2. The use of Padé approximations

Our point of departure is the use of Padé approximations (with linear numerators) to $f(x)$ of the form

$$u(x)=\frac{x+\gamma}{P_{r-2}(x)}, \quad r\geqslant 2. \tag{1}$$

Here $P_{r-2}(x)$ is a polynomial of degree $(r-2)$ whose $(r-1)$ coefficients, along with γ, are subjected to the conditions

$$u^{(k)}(x_n)=f^{(k)}(x_n), \quad k=0,1,\dots,r-1, \tag{2}$$

where x_n is the nth approximation to the required root $x=\rho$. We restrict our attention to an interval about $x=\rho$ for which (1)–(2) possess a unique solution. This 'locality' restriction is to be expected in higher order methods. Even for the Newton–Raphson method $(r=2)$ $f'(x)\neq 0$ is needed in the relevant interval. The coefficients of P_{r-2}, as well as γ, will be functions of x_n and $f^{(k)}(x_n)$, $k=0,1,\dots,r-1$. The root of $u(x)$, i.e., $(-\gamma)$, is taken to be the next approximation x_{n+1}, leading to the next Padé approximation u. Thus,

$$-\gamma=x_{n+1}=x_n-P_{r-2}(x_n)f(x_n). \tag{3}$$

Henceforth we shall suppress the index $(r-2)$ and write P instead of P_{r-2}.

Since we are only interested in $P(x_n)$ and *not* in its individual coefficients, we might use (1) to differentiate $P(x)\,u(x)$ and then invoke (2) to arrive at the system

$$(Pf)'=1, \qquad \text{at } x=x_n, \tag{4a}$$

$$(Pf)^{(k)}=0, \quad k=2,3,\dots,r-1, \quad \text{at } x=x_n. \tag{4b}$$

However, the attempt to solve the system (4) merely for the unknown $P(x_n)$, leads to a quotient of rather cumbersome determinants not obviously expressible in closed form. Similar determinants, for small values of r, were displayed in the formulas of Domoryad [10, p. 113]. Since our goal is a closed form iterative formula, we shall resort to a different approach.

Let us rewrite (1) in the form

$$P(x)=(x+\gamma)(1/u(x)). \tag{5}$$

Differentiating (5) m times, we get

$$\begin{aligned} P^{(m)}&=\sum_{k=0}^{m}\binom{m}{k}(x+\gamma)^{(k)}\left(\frac{1}{u}\right)^{(m-k)}\\ &=(x+\gamma)\left(\frac{1}{u}\right)^{(m)}+m\left(\frac{1}{u}\right)^{(m-1)}=Pu\left(\frac{1}{u}\right)^{(m)}+m\left(\frac{1}{u}\right)^{(m-1)}, \end{aligned} \tag{6}$$

where $(x+\gamma)$ has been replaced by Pu in the last member of (6). But $P^{(r-1)}=0$, since P is of degree $(r-2)$, and thus for $m=r-1$ we obtain from (6)

$$Pu=\frac{(1-r)(1/u)^{(r-2)}}{(1/u)^{(r-1)}}. \tag{7}$$

Evaluating (7) at $x=x_n$ and using (2), we reach

$$(Pf)_n=\frac{(1-r)(1/f)_n^{(r-2)}}{(1/f)_n^{(r-1)}}, \tag{8}$$

where the subscript n indicates that the relevant quantity is evaluated at $x = x_n$. Finally we substitute (8) in (3) to obtain the iterative formula

$$x_{n+1} = x_n - (1-r)\frac{(1/f)_n^{(r-2)}}{(1/f)_n^{(r-1)}}. \tag{9}$$

This formula ties in with Theorem 4.4.2 of Householder [5, p. 169], and can readily be shown to be of rth order.

As stated, our objective is to *view* (9) *as describing Newton–Raphson iterations applied to an as yet unknown function* $g(x)$, *having the same simple root* $x = \rho$ *as* $f(x)$. Thus we rewrite (9) in the form

$$x_{n+1} = x_n - g_n/g_n', \tag{10}$$

so that we must have, for every x,

$$\frac{g'}{g} = \frac{-1}{r-1}\,\frac{(1/f)^{(r-1)}}{(1/f)^{(r-2)}}. \tag{11}$$

Integration of (11) leads to

$$g = \left[(C/f)^{(r-2)}\right]^{-1/(r-1)}, \tag{12}$$

where C is an arbitrary constant. Since, at this point, we are only interested in the computationally relevant quantity g/g', the value of C is *immaterial* and we may choose its value, for simplicity, at will. It can be shown, using Faa de Bruno's formula [6, pp. 33–34], that

$$g = \frac{Kf}{\left[(f')^{r-2} + \sum_{j=1}^{r-3}(-1)^j f^j \varphi_j\right]^{1/(r-1)}}. \tag{13}$$

Here K is a constant depending only upon r, while φ_j is a homogeneous polynomial, with positive coefficients, of the derivatives of f up to and including order $(j+1)$. For example, $\varphi_1 = \frac{1}{2}(r-3)(f')^{r-4}f''$, and we may calculate additional φ_j's as well. However, the exact form of the φ_j's will not be needed and is therefore omitted. Explicit forms of g as well as g/g', for a few values of r, will be given in the next section.

Since f was assumed sufficiently smooth, it is clear that so will be g. In particular, g and its derivatives are finite at $x = \rho$ by (13) and the fact that $f'(\rho) \neq 0$.

It follows from (13) that

$$g(\rho) = f(\rho) = 0, \tag{14}$$

and therefore,

$$\lim_{x\to\rho}\frac{g'(x)}{f'(x)} = \lim_{x\to\rho}\frac{g(x)}{f(x)} = K\left[f'(\rho)\right]^{-(r-2)/(r-1)}. \tag{15}$$

Without loss of generality we may assume $f'(\rho) > 0$, and we see from (15) that K may be chosen, in principle, so that

$$g'(\rho) = f'(\rho) > 0. \tag{16}$$

Regardless of the choice of $K \neq 0$, however, we see from (15) that g, as well as f, has a simple root at $x = \rho$.

3. Some explicit forms of g

We shall now examine the explicit forms of g for $r = 2, 3, 4, 5$. The case $r = 2$ is trivial, and it follows from (12) with $C = 1$ that $g = f$. Thus (9) and (10) reduce to the standard Newton–Raphson iterations.

For $r = 3$ we draw from (12), (13) that

$$g = \frac{Kf}{(f')^{1/2}}, \quad r = 3, \tag{17}$$

and (10) takes the form

$$x_{n+1} = x_n - \frac{f_n f_n'}{(f_n')^2 - \frac{1}{2} f_n f_n''}. \tag{18}$$

This is the well-known third order Halley formula [4], [9, p. 91] also known as the method of tangent hyperbolas. The case $r = 3$ has been treated by Jordan [7] who pointed out that "any operation that reduces the curvature will improve the convergence of the iteration". Indeed,

$$g'' = Kf \frac{3(f'')^2 - 2f'f'''}{4(f')^{5/2}}, \quad r = 3, \tag{19}$$

so that $g''(\rho) = 0$. The significance of this fact will be further explored in the next section. Some global convergence properties of (18) have been discussed in [2].

For $r = 4, 5$ we obtain

$$g = \frac{Kf}{\left[(f')^2 - \frac{1}{2} ff''\right]^{1/3}}, \quad r = 4, \tag{20}$$

$$g = \frac{Kf}{\left[(f')^3 - ff'f'' + \frac{1}{6} f^2 f'''\right]^{1/4}}, \quad r = 5. \tag{21}$$

Formulas (17), (20) and (21) are of course special cases of (13). Introduction of (20) and (21) into (10) leads to

$$x_{n+1} = x_n - \frac{f_n\left[(f')^2 - \frac{1}{2} ff''\right]_n}{\left[(f')^3 - ff'f'' + \frac{1}{6} f^2 f'''\right]_n}, \tag{22}$$

$$x_{n+1} = x_n - \frac{f_n\left[(f')^3 - ff'f'' + \frac{1}{6} f^2 f'''\right]_n}{\left[(f')^4 - \frac{3}{2} f (f')^2 f'' + \frac{1}{4} f^2 (f')^2 + \frac{1}{3} f^2 f' f''' - \frac{1}{24} f^3 f''''\right]_n} \tag{23}$$

for $r = 4, 5$, respectively. Formulas (22) and (23) have been obtained by Kiss [8], and are of order four and five respectively. Corresponding iterative formulas for extraction of kth roots, using $f(x) = x^k - a$, have been derived by Breuer and Zwas [1] for $r \leqslant 5$ via *elementary algebra.*

Ehrman [3] analyzed the relative computational merits of higher order iterative methods, containing higher derivatives. Computational efficiency, however, is not our concern here. Our point of view is to regard an rth order iteration as being derived from a rectified function g, whose intrinsic approximation properties will be studied in the next section.

4. Properties of g

Standard analysis shows that (9), and hence (10), are iterations of rth order, i.e.,

$$x_{n+1} - \rho = \mathrm{O}\left[(x_n - \rho)^r\right], \tag{24}$$

and we shall proceed to demonstrate intrinsic properties of g justifying its description as a rectified approximation to $f(x)$ about ρ. *We shall prove that* $g^{(k)}(\rho) = 0$ *for* $k = 2, 3,\dots, r-1$. To this end, consider the Taylor expansion of $g(\rho) = 0$ about x_n,

$$g(\rho) = g(x_n) + g'(x_n)(\rho - x_n) + \sum_{k=2}^{r-1} \frac{g^{(k)}(x_n)}{k!}(\rho - x_n)^k + \mathrm{O}\left[(\rho - x_n)^r\right], \tag{25}$$

and the Newton–Raphson iteration applied to g,

$$0 = g(x_n) + g'(x_n)(x_{n+1} - x_n). \tag{26}$$

Subtracting (26) from (25) and using $g(\rho) = 0$, we reach

$$g'(x_n)(\rho - x_{n+1}) + \sum_{k=2}^{r-1} \frac{g^{(k)}(x_n)}{k!}(\rho - x_n)^k = \mathrm{O}\left[(\rho - x_n)^r\right]. \tag{27}$$

Invoking (24), we find from (27) that

$$\sum_{k=2}^{r-1} \frac{g^{(k)}(x_n)}{k!}(\rho - x_n)^k = \mathrm{O}\left[(\rho - x_n)^r\right]. \tag{28}$$

Next we expand $g^{(k)}(x_n)$ about ρ up to powers of $(r-k-1)$ and substitute into (28), to obtain

$$\sum_{k=2}^{r-1} \frac{(\rho - x_n)^k}{k!}\left[\sum_{j=0}^{r-k-1} \frac{g^{(k+j)}(\rho)}{j!}(x_n - \rho)^j\right] = \mathrm{O}\left[(x_n - \rho)^r\right]. \tag{29}$$

Setting $k + j = l$, (29) can be rewritten as

$$\sum_{k=2}^{r-1} \sum_{l=k}^{r-1} \frac{(-1)^k}{k!(l-k)!} g^{(l)}(\rho)(x_n - \rho)^l = \mathrm{O}\left[(x_n - \rho)^r\right]. \tag{30}$$

Interchanging the order of summation and multiplying up and down by $l!$, we find that

$$\sum_{l=2}^{r-1}\left[\sum_{k=2}^{l} \binom{l}{k}(-1)^k\right] \frac{g^{(l)}(\rho)(x_n - \rho)^l}{l!} = \mathrm{O}\left[(x_n - \rho)^r\right]. \tag{31}$$

The inner summation is readily seen to equal $(l-1)$, so that we have

$$\sum_{l=2}^{r-1} \frac{g^{(l)}(\rho)}{l(l-2)!}(x_n-\rho)^l = \mathrm{O}\left[(x_n-\rho)^r\right]. \tag{32}$$

Dividing both members of (32) by $(x_n-\rho)^2$ and letting x_n tend to ρ, we find that $g''(\rho)=0$ for every $r \geqslant 3$. Similarly, we divide by $(x_n-\rho)^k$ for $k=3, 4,\dots,r-1$, let x_n tend to ρ, and find that $g^{(k)}(\rho)=0$ for the same values of k. This is what we set out to prove.

Collecting our results, including (14) and (16), we find the function g to possess the following properties:

$$g(\rho)=0, \qquad g'(\rho)\neq 0, \qquad g^{(k)}(\rho)=0, \quad k=2, 3,\dots,r-1. \tag{33}$$

It is those properties which make g a rectification of f about ρ, since it evidently has an $(r-1)$st order contact with the tangent line at ρ. The Newton–Raphson method (10) applied to g, therefore, naturally produces rth order iterations as seen from (24).

Since $f(\rho)=0$ and $f'(\rho)\neq 0$, we have $f=\mathrm{O}(x_n-\rho)$, and (33) shows that for $k=2, 3,\dots,r-1$,

$$g^{(k)} = \mathrm{O}\left[(x_n-\rho)^{r-k}\right] = \mathrm{O}(f^{r-k}), \tag{34}$$

in the neighborhood of ρ as exemplified in (19) for $r=3$, $k=2$. In particular,

$$g'' = \mathrm{O}(f^{r-2}), \tag{35}$$

in the neighborhood of ρ, a property which will be made use of in the last section.

5. Recursive relations

Various recursive relations among g's of successive orders and their derivatives are derivable from (12). For example, it follows immediately from (12) that, for arbitrary C,

$$\left(\frac{1}{g_{r+1}}\right)^r = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{g_r}\right)^{r-1}, \tag{36}$$

where, throughout this section, g_r denotes the rectified approximation of order r. If we choose, for simplicity, $C=C_r=(-1)^r/(r-2)!$ in the definition (12) of g_r, (36) will take the form

$$g_{r+1} = g_r(g_r')^{-1/r}. \tag{37}$$

Now we set $h_r=(g_r')^{-1/r}$ and observe that all derivatives of h_r^α, for any α, are finite at $x=\rho$ by (16). Differentiating (37) k times, we have

$$g_{r+1}^{(k)} = \sum_{j=0}^{k}\binom{k}{j} g_r^{(j)} h_r^{(k-j)}. \tag{38}$$

Suppose we know that

$$g_r''(\rho) = g_r'''(\rho) = \cdots = g_r^{(r-1)}(\rho) = 0, \tag{39}$$

and set $x=\rho$ in (38). For $2 \leqslant k \leqslant r$ this yields

$$g_{r+1}^{(k)} = k g_r' h_r^{(k-1)} + g_r^{(k)} h_r, \quad x=\rho, \tag{40}$$

since $g_r(\rho)=0$. For $2 \leqslant k \leqslant r-1$ the second term on the right of (40) vanishes by (39) while

$$h_r^{(k-1)} = \left[(g_r')^{-1/r}\right]^{(k-1)} = \frac{-1}{r}\left[h_r^{r+1} g_r''\right]^{(k-2)}$$
$$= \frac{-1}{r}\sum_{j=0}^{k-2}\binom{k-2}{j} g_r^{(j+2)}\left(h_r^{r+1}\right)^{(k-2-j)} = 0, \quad x=\rho, \tag{41}$$

again by (39). On the other hand, for $k=r$, (40) becomes

$$g_{r+1}^{(r)} = r g_r' h_r^{(r-1)} + g_r^{(r)} h_r, \quad x=\rho, \tag{42}$$

and setting $k=r$ in (41), we obtain

$$h_r^{(r-1)} = \frac{-1}{r} g_r^{(r)} h_r^{r+1}, \quad x=\rho. \tag{43}$$

Recalling that $g_r' = h_r^{-r}$ and introducing (43) into (42) we find $g_{r+1}^{(r)}(\rho)=0$. Collecting our results, we conclude that (39) implies

$$g_{r+1}''(\rho) = g_{r+1}'''(\rho) = \cdots = g_{r+1}^{(r)}(\rho) = 0. \tag{44}$$

Recalling that $g_3''(\rho)=0$ from (19), we have arrived at an inductive proof of the rectification properties (33) following directly from the recursive relations (36).

Clearly, the rectified approximation g_r of f about ρ is non-unique, since adding to it terms of order $(x-\rho)^r$ will leave its rectification properties conserved. In particular, $\tilde{g}_r = g_r(g_r')^{\alpha}$ for any α, is a family of rectified approximations, which is readily seen to satisfy (33).

6. Final remarks

At this point some remarks about the root finding iterations (10), based upon g, are in order. Investigation of their convergence is much facilitated by the use of g since, once g has been constructed, the extensive available theory of Newton–Raphson iterations can be applied. A typical result of this theory ensures convergence if for every x in the relevant interval,

$$\left|\frac{f(x)f''(x)}{f'(x)^2}\right| \leqslant \lambda < 1. \tag{45}$$

We note that near a simple root $x=\rho$,

$$\left|\frac{f\cdot f''}{(f')^2}\right| = O(f), \tag{46}$$

whereas by (13) and (35) the same result applied to g takes the form

$$\left|\frac{g\cdot g''}{(g')^2}\right| = O(f^{r-1}). \tag{47}$$

Consequently, it stands to reason that in many cases condition (45), with f replaced by g, is less severe than (45) itself, for $r \geqslant 3$.

In order to illustrate this phenomenon, consider $f = xe^{-x/10}$ which has a simple root at $x = 0$. Then,

$$\frac{ff''}{(f')^2} = \frac{x(x-20)}{(10-x)^2}, \tag{48}$$

and we readily find that (45) is satisfied for

$$a < x < 5(2-\sqrt{2}) \approx 2.929, \tag{49}$$

where a is any finite negative number. On the other hand, with $r = 3$, we have $g = xe^{-5x}(1-10x)^{-1/2}$ and we find

$$\frac{gg''}{(g')^2} = \frac{50x^2(50x^2-20x+3)}{(50x^2-10x+1)^2}. \tag{50}$$

This time (45), with f replaced by g, is seen to hold for

$$a < x < 10(2-\sqrt{2}) \approx 5.858, \tag{51}$$

thus doubling the positive interval in which convergence is assured by (45).

Finally, for the case of a multiple root of f, we routinely replace f by f/f' so that (12) becomes

$$g = \left[(Cf'/f)^{(r-2)}\right]^{-1/(r-1)}, \tag{52}$$

which is to be introduced in (10).

It is intriguing to investigate the extension of the rectification method to systems of nonlinear equations. Results along these lines will be published in a future paper.

References

[1] S. Breuer and G. Zwas, Computer root extraction by a-priori design, *Comput. Education,* to appear.

[2] M. Davies and B. Dawson, On the global convergence of Halley's iterative formula, *Numer. Math.* **24** (1975) 133–135.

[3] H. Ehrman, Konstruktion und Durchführung von Iterationsverfahren höherer Ordnung, *Arch. Rat. Mech. Anal.* **4** (1960) 65–88.

[4] E. Halley, A new and general method of finding the roots of equations, *Philos. Trans. Roy. Soc.* **18** (1796) 136.

[5] A.S. Householder, *The Numerical Treatment of a Single Nonlinear Equation* (McGraw-Hill, New York, 1970).

[6] C. Jordan *Calculus of Finite Differences* (Chelsea, New York, 1965).

[7] W.B. Jordan, An iterative process, *Mathematical Tables and Aids to Computation* **5** (1951) 183.

[8] I. Kiss. Über eine Verallgemeinerung des Newtonschen Näherungsverfahrens, *Z. Angew. Math. Mech.* **34** (1954) 68–69.

[9] J.F. Traub, *Iterative Methods for the Solution of Equations* (Prentice-Hall, Englewood Cliffs, NJ, 1964).

[10] V.L. Zaguskin, *Handbook of Numerical Methods for the Solution of Algebraic and Transcendental Equations* (Pergamon Press, Oxford, 1961).